基于深度学习的人脸识别算法研究 *

胡亚洲，周亚丽，张奇志

(北京信息科技大学 自动化学院，北京 100192)

摘　要：基于深度学习的人脸识别技术是目前人工智能领域研究的热点之一。考虑到现实环境中的人脸图片在角度、光线、分辨率上的复杂程度，对 Inception-ResNet-V1 网络结构进行了改进，同时完成了数据集制作、超参数调节等相关工作，并在家庭服务机器人平台上进行了实验研究。实验结果表明，改进的网络结构在 LFW 测试集上准确率达到 99.22%，高于原始网络结构的 99.05%；在亚洲人脸数据集上准确率达到 99.20%，高于原始网络结构的 97.10%；在自建非匹配人脸数据集上误识别率为 3.43%，低于原始网络结构的 12.28%。可以看出，和原始网络结构相比，改进的网络结构提升了人脸识别的准确率和降低了误识别率。

关键词：家庭服务机器人；人脸识别；深度学习；Inception-ResNet-V1

中图分类号：TP391.4　　doi: 10.19734/j.issn.1001-3695.2018.09.0815

Face recognition algorithm based on deep learning

Hu Yazhou, Zhou Yali, Zhang Qizhi

(*School of automation, Beijing Information Science & Technology University, Beijing* 100192, *China*)

Abstract: Face recognition technology based on deep learning is one of the hot topics in the field of artificial intelligence. Considering the complexity of the angle, light and resolution of face images in real environment, this paper improves the network structure of Inception-ResNet-V1, completes the related work of dataset production and hyper-parameter adjustment, and carries out experimental research on the home service robot platform. The experimental results show that the improved network in this paper achieve 99.22% accuracy in LFW testset, which is higher than 99.05% of the original network structure. It reaches 99.20% accuracy on the Asian face dataset, which is 97.10% higher than the original network structure. The false recognition rate on self built mismatched face dataset is 3.43%, which is lower than 12.28% of the original network structure. It can be seen that compared with the original network structure, the improved network structure improves the accuracy of face recognition and reduces the false recognition rate.

Key words: home service robot; face recognition; deep learning; Inception-ResNet-V1

0　引言

家庭服务机器人[1]是指装备有视觉、罗盘、激光雷达等传感器，是为人类服务的特种机器人，能够代替人完成家庭服务工作的机器人。

人脸识别是基于人的脸部特征信息进行身份识别的一种生物识别技术[2]。用摄像头或者抓拍机采集包含人脸的图像或视频流，并通过相关的算法自动在图像或视频流中检测人脸，获取人脸的位置信息，再对人脸进行特征提取、特征比对、输出人脸信息等一系列相关技术，通常也叫做人像识别或面部识别。人脸识别是当前人工智能领域研究的热点之一，众多的学者和科研工作者不断地优化现有算法或者提出新的算法，以提高人脸检测与识别的速度与准确度及降低误检率。

随着人工智能的发展，人脸识别技术和家庭服务机器人的研究都得到一定程度的发展，把人脸识别技术应用到家庭服务机器人中，也是计算机视觉和机器人领域研究的热点之一。

在家庭服务机器人研究比较发达的国家，如加拿大、日本、美国等，各类家庭服务机器人和工业机器人研究大多是企业和科研院所共同承担的，将研究、产业发展和市场需求非常好的结合起来，使研究服务于企业，企业促使研究发展，形成了一个良性循环。在学术界和工业界都具有良好的开发环境，故而这些国家的家庭服务机器人目前已经取得了较好的研究成果。

2008 年，加拿大推出了完美机器人妻子 Aiko，该女性机器人一头秀发、五官精致，能够从事清洁和家务工作、善于数学计算、可以识别多个家庭成员、能够朗读报纸、为他人指引方向等[3]。

2010 年 5 月，日本 Fujitsu（富士通）公司推出了一款家用医疗康复机器人泰迪熊，该机器人能够快速识别人脸面部表情，并作出相应反映。泰迪熊身上的传感器可以收集使用者情绪信息和感知周围环境，并找到适当的方式和使用者互动[3]。

2017 年 5 月，全球领先的美国家用机器人公司 iRobot 宣布在中国市场推出全新的扫地机器人。iRobot 研发出的扫地机器人 Roomba 可以有效地清除灰尘。iRobot 独创了 Vslam 图像位移定位系统，利用摄像头自动定位和地图构建，一切数据的获取，都是通过设备顶部的摄像头来完成，但摄像头的工作误差，会直接影响定位的精准，进而影响机器人的工作效率。

收稿日期：2018-09-17；**修回日期：**2018-11-28　**基金项目：**国家自然科学基金资助项目（11672044，11172047）；北京信息科技大学教改项目（2016JGYB09）

作者简介：胡亚洲（1993-），男，河北邯郸人，硕士研究生，主要研究方向为深度学习图像处理（18800110029@163.com）.；周亚丽（1968-），女，辽宁沈阳人，教授，博士，主要研究方向为机器人控制及信号处理究；张奇志（1963-），男，辽宁阜新人，教授，博士，主要研究方向为机器人控制.

我国的家庭服务机器人起步较晚，与国外的服务机器人的发展水平还有一定差距，但是广阔的应用前景还是吸引着越来越多的科研工作者投入到家庭服务机器人的研究领域来。其中，比较有代表性的是中国科学技术大学自主研发的智能服务机器人“可佳”[4]。该机器人具有自然语言人机交互、自动推理与知识获取、环境感知与建模、机器人控制等核心技术。

Sun@Home[5]是北京信息科技大学的一支研发家庭服务机器人的团队。自 2010 年以来，该团队研发的家庭服务机器人参加国内外机器人竞赛和公益展示十余场，取得了骄人的成绩，获得国内外同行的好评。在目标分割、室内定位与导航、移动取物、声源定位与语音识别、物体分类识别、人脸识别等关键技术有深入的研究，部分技术已达到国内领先水平。

人脸识别是一个典型的图像模式分析、理解与分类计算问题，涉及多个交叉学科。1965 年 Chan 等人[6]在 Panoranic Research Inc 发表了技术报告，从此揭开了人脸识别研究的序幕。近二、三十年来，越来越多的科研院所和互联网公司投入到人脸识别的研发中，并且在人脸识别算法研究和应用创新取得了非常瞩目的成绩。例如 1991 年，美国麻省理工学院（MIT）由 Pentland 领导的研究小组提出的具有里程碑意义的特征脸[7]方法；1996 年由 Kriegman 领导的研究小组则提出了同样具有重大意义的 FisherFace[8]方法；1997 年，清华大学彭辉等人[9]对特征脸的方法做了进一步的改进，提出采用类间散布矩阵作为产生矩阵，该方法降低了产生矩阵的维数，进而提高了特征计算的速度，而且人脸识别的准确率不受影响；2005 年，密西根州立大学（MSU）由 Jain 领衔的小组在 3-D 人脸识别[10]方法的研究做了大量工作，取得了较为突出的效果。传统算法为人脸识别的研究提供了良好的基础。

近几年，由于深度学习的广泛应用，在人脸识别上有了很大的突破。2014 年，在 IEEE 国际计算机视觉与模式识别会议(IEEE Conference on Computer Vision and Pattern Recognition, CVPR)上，Facebook 提出了基于深度学习的人脸识别 DeepFace[11]算法，在 Labeled Faces in the Wild(LFW)上取得了 97.35%的识别率，接近人用肉眼在 LFW 上的识别率 97.52%。该算法利用卷积网络预测输出向量，将最高的隐含层作为人脸特征，这一层在训练过程中要区分大量的人脸类别，因此包含了丰富的类间变化的信息，有很强的泛化能力。

2014 年 6 月，汤晓鸥带领　　中文大学计算机视觉研究组开发了一个名为 DeepID [12,13]的深度学习人脸识别算法，在 LFW 数据库上获得了 99.15%的识别率。在此之前，汤晓鸥的研究组开发了一个基于高斯过程的人脸识别技术 GaussianFace，取得了 98.52%的识别率。这也是计算机自动识别算法的识别率首次超过人眼。

2015 年，Google 提出基于深度学习的人脸识别算法 FaceNet[14]。该算法与其他的深度学习方法在人脸上的应用不同，FaceNet 没有用传统的 Softmax 的方式去进行分类学习，然后抽取其中某一层作为特征，而是直接进行端对端学习一个从图像到欧氏空间的编码方法，输出人脸特征向量，然后基于这个特征向量再进行人脸识别、人脸验证和人脸聚类等。

2017 年，Liu 等人[15]提出 SphereFace 深度学习人脸识别算法，该算法主要提出了归一化权值和角度间距，基于这两点，对传统的 Softmax 进行了改进，从而实现了最大类内距离小于最小类间距离的识别标准。该算法识别率准确率高，在小数据集(少于 50W 的训练数据)上一直保持验证准确率第一。

2018 年，腾讯提出基于大间隔余弦损失的深度学习人脸识别算法 CosFace[16]。该算法主要对人脸识别的损失函数进行了优化，使用余弦距离作为损失函数。余弦距离更多的是从方向上区分差异，在人脸识别上具有很好的泛化能力。该算法表现优异，不仅在 LFW 取得很高的识别率，而且将 MegaFace: million-scale face recognition(MegaFace)数据集的精度提升到 98%，超过俄罗斯 Vocord 公司保持的 91% 的纪录。

人脸识别主要步骤为人脸检测、人脸裁剪、特征提取、特征比对得出余弦距离，根据阈值判断是否为同一个人。人脸识别算法实现的流程如图 1 所示。本文中的人脸识别是在家庭服务机器人上应用，由于家庭环境场景极为复杂，容易受到光线、角度、人脸大小等影响。传统的人脸识别算法，如 LBP 人脸识别、FisherFace 等算法，鲁棒性差，导致在家庭环境中识别效果很差，误识别严重，因此不能很好地应用在家庭服务机器人上。

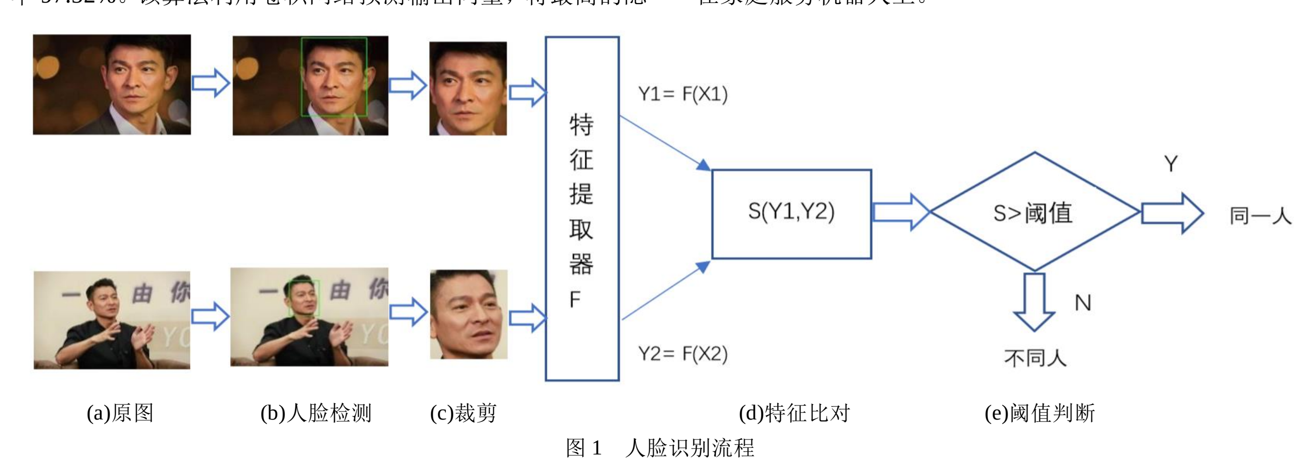


图 1　人脸识别流程

Fig. 1　Flow chart of face recognition

深度学习卷积神经网络人脸识别[17]具有很好的鲁棒性，适应多场景的识别，在一定程度上可以有效减少因为光线、角度等导致的识别率下降。但是目前衡量人脸识别准确率的主要指标是 LFW 的准确率，深度学习人脸识别在 LFW 的准确率都很高，但是在服务机器人上测试人脸识别相关模型，准确率下降很大，不能完全满足现有需求,主要原因是现实环境中的人脸图片在角度、光线、分辨率上复杂程度远远高于 LFW 数据集。

针对上述问题分析，结合家庭服务机器人的具体应用场景，本文提出改进的 Inception-ResNet-V1 网络结构。该网络

结构用在人脸特征向量提取，同时完成了亚洲人脸数据训练、超参数调节等相关工作。通过与 FaceNet 比较，本文的实验结果在 LFW 数据集和实际采集的家庭环境人脸数据集上均优于原始网络结构。

1 算法实现

1.1 Inception-Resnet-V1 网络结构

Inception-Resnet-V1 网络结构如图 2 所示。该网络是

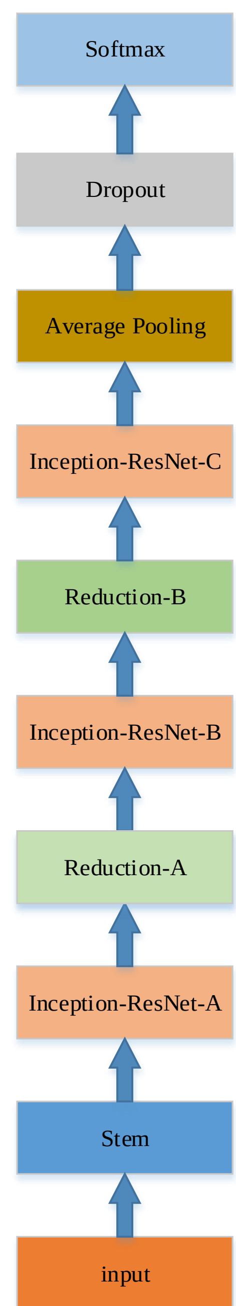


图 2 Inception-ResNet-V1 网络结构

Fig. 2 Network structure of Inception-ReSnet-V1

Inceptipn 网络与 ResNet 网络的结合[18]，网络包含 6 个卷积层的主干网络(Stem)，5 层的 Inception-ResNet-A、Reduction-A 网络，10 层的 Inception-ResNet-B、Reduction-B 网络，5 层的 Inception-ResNet-C 网络、池化层(Average Pooling)、Dropout 层和 Softmax 分类器。该网络结构结合了 Inception 和 ResNet 网络的优点，不仅增加了网络的宽度，使网络的适应性更强，泛化能力显著加强，同时加深网络的深度，提升网络特征提取能力，进而提高分类的准确率。其中，Inception-ResNet-A、B、C 是 Inception 网络与 ResNet 网络的组合结构。池化层是对输入进行降采样，降低输出数据的维度，并且不损失网络提取的显著特征。Dropout 层是一种深度学习正则化方法，可以减少过拟合。Softmax 分类器对多个类别进行分类，将多个神经元的输出，映射到（0,1）区间内，从而来进行多分类。

网络结构如果使用同样的卷积核，会造成特征单一，丢失部分有用特征，导致特征提取能力减弱。如果使用不同的卷积核，可以获取不同大小的感受野，多尺度特征融合，提高模型泛化能力。

Inception 的网络，将 1×1、3×3、5×5 的卷积核和 3×3 的池化层堆叠在一起，一方面增加了网络的宽度，另一方面增加了网络对尺度的适应性。Inception 网络结构如图 3 所示。

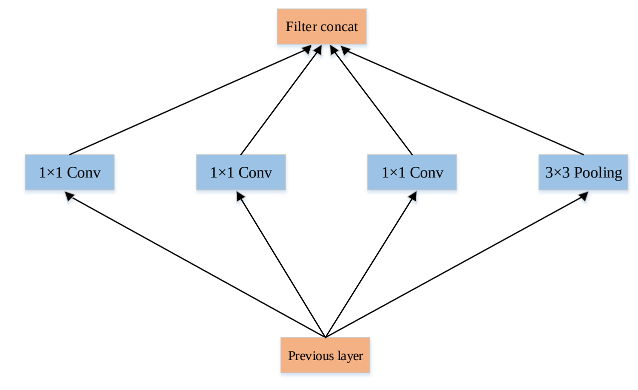


图 3 Inception 网络结构

Fig. 3 Network structure of Inception

GoogLeNet[19]是 Inception 的代表作，显示了网络有足够的深度是模型表现良好的前提，但是在网络达到一定深度之后，简单的网络堆叠反而效果变差。而且随着网络结构深度的增加，计算参数增加，导致网络性能下降。在网络深度到一定程度时，更深的网络意味着更高的训练误差。误差升高的原因是网络越深，梯度消失的现象就越明显。

ResNet[18]网络是在 2015 年提出，获得 ImageNet 比赛分类任务第一名，该网络结构的提出就是解决在增加网络深度的情况下可以有效解决梯度消失的问题。ResNet 中核心结构是残差网络，如图 4 所示：

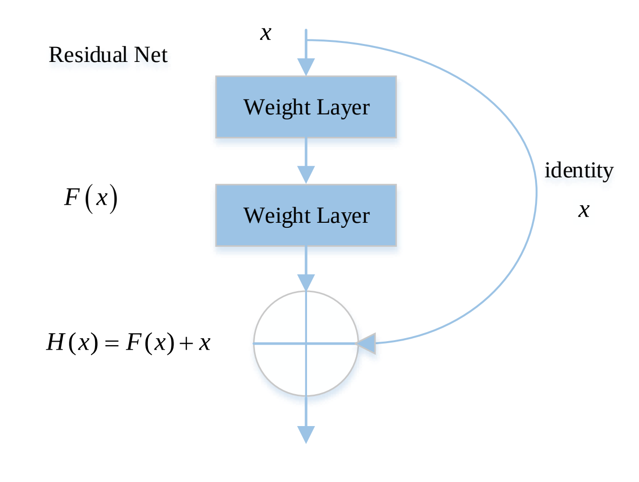


图 4 残差网络示意图

Fig. 4 Network structure of Resnet

Residual Net：$H(x)$ 是输出，希望 2 层权重能够拟合 $F(x)$，使得 $F(x)$ 是一个关于恒等式残差的映射。x 是 identity mapping，为上一层网络的输出。Residual Net 的公式为

$$H(x) = F(x) + x \quad (1)$$

为增强网络的非线性映射能力，同时限制网络规模的大

小，网络在提取特征后，接入一个全连接层。该层的每一个神经元与前一层的所有神经元互相连接，同层神经元之间不连接。公式为

$$o_j^l = f(\sum_{i=1}^{n} x_i^{l-1} w_{ji}^{(l)} + b_j^{(l)}) \tag{2}$$

其中：n 是前一层的神经元个数；l 表示当前层数；$w_{ji}^{(l)}$ 是该层神经元 j 与前一层神经元 i 的连接强度；$b_j^{(l)}$ 是该层神经元 j 的偏置；$f(\bullet)$ 表示激活函数。

1.2 Inception-Resnet-V1 网络结构改进

本文采用的一个数据集是 CASIA -webface，该数据集有 10 575 个 ID（同一个人所有人脸图像存在同一个文件夹下，称为一个 ID）和 494 414 张图片。该数据集存在人脸图像少、分布不均衡、噪声多的问题，使用 FaceNet 官方的网络结构，在 LFW 上测试准确率为 99.05%。通过研究和测试该数据集发现，网络的结构对最终的识别率有很大的影响。因此本文在 Stem 主干网络下添加了改进的 Inception 网络模块，该模块具有增加网络深度和宽度，获取更多显著特征的作用。网络结构如图 5 所示。同时修改 Inception-ResNet-A、Inception-ResNet-B、Inception-ResNet-C 三个子网络的层数和 Inception 层中的子网络权重参数，把 Inception-ResNet 子网络层数修改为 20 层，可以提取更为有效的特征数据。在 Inception-ResNet-C 输出层添加卷积核 1×1 的卷积降维层，该卷积层起到一个跨通道聚合的作用，在保证准确率的前提下有效地减少参数，提升运算速度和减小模型大小。

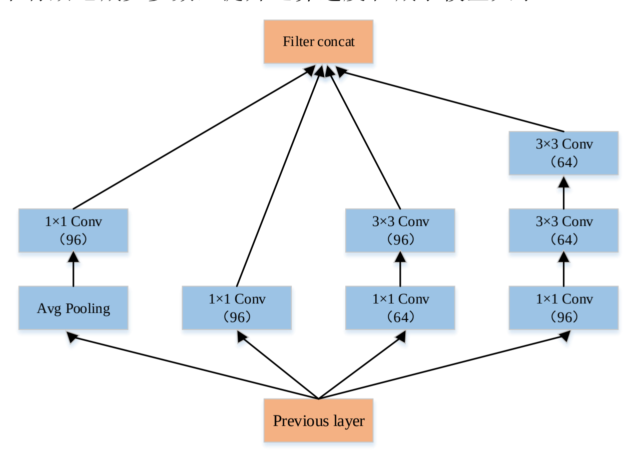


图 5 添加的 Inception 模块示意图

Fig. 5 Network structure of Inception

原始的 Inception-ResNet-V1 网络在 Stem 主干网络后直接连接 Inception-ResNet 网络组合模块，Inception 模块中参数比例系数为 0.2，这种网络模型在结构上丢失了一部分特征信息，网络不能完全获取人脸的有用特征，进而会影响人脸识别的准确率。因此，在 Stem 主干网络下添加多通道多卷积核的 Inception 层，该层通过不同的卷积核参数，获取不同感受野的特征，然后聚合不同感受野的特征，有效地改进原始网络特征提取不足的问题。模块中添加的 1×1 卷积层，只有一个卷积核，为特征图增加了一个尺度变换，对识别精度有所提高。该结构可以同时增加了网络的深度和宽度，可以提取到更高、更抽象的人脸特征，加强了人脸识别的泛化能力。

本文的网络结构用在人脸特征提取，这是决定人脸识别准确率的关键步骤，特征提取不准确，导致计算欧氏距离或者余弦距离的结果会有偏差。本文中使用余弦距离作为特征比对算法。欧氏距离是计算两个向量间的绝对距离，越接近 0 表示两个向量越相似，但是存在同一个人的两张不同图片，由于光照、角度等因素导致欧氏距离很大，这样使得系统判定为两个人脸不是同一个人，导致比对错误。N 维空间 $a(x_{11}, x_{12}, \cdots, x_{1n})$ 点与 $b(y_{11}, y_{12}, \cdots, y_{1n})$ 点间的欧氏距离为 d_{12}。欧氏距离公式为

$$d_{12} = \sqrt{\sum_{k=1}^{n} (x_{1k} - x_{2k})^2} \tag{3}$$

余弦距离是用向量空间中两个向量夹角的余弦值作为衡量两个个体间差异的大小，余弦相似度更加注重两个向量在方向上的差异，而非距离或长度上。在实验中，保持测试数据集、参数、网络结构等所有前置条件不变，使用余弦距离作为特征比对算法的准确率优于欧氏距离。N 维空间 $a(x_{11}, x_{12}, \cdots, x_{1n})$ 点与 $b(y_{11}, y_{12}, \cdots, y_{1n})$ 点间的余弦距离为 $\cos\theta$。余弦距离公式为

$$\cos\theta = \frac{\sum_{k=1}^{n} x_{1k} x_{2k}}{\sqrt{\sum_{k=1}^{n} x_{1k}^2} \sqrt{\sum_{k=1}^{n} x_{2k}^2}} \tag{4}$$

2 实验研究

2.1 实验平台

图 6 是北京信息科技大学家庭服务机器人团队的 Sun@Home 机器人。



图 6 Sun@Home 家庭服务机器人

Fig. 6 Home service robot of Sun@Home

本实验平台主要由一个 Kinect-v2 摄像头[20]、可调升降机构、3 自由度机械臂、全向轮底盘和一个检测范围为 270° 的激光雷达组成。本次实验使用到的传感器装置是 Kinect-v2 摄像头。相比于 2010 年发布第一代 Kinect-v1，第二代 Kinect 感应器可以获取最大摄像头分辨率为 1920×1080，高于 Kinect-v1 版本的 640×480；最高帧率支持 30 fps；检测范围为 0.5~4.5 m；检测角度水平为 70°，垂直为 60°。各项性能指标均优于 Kinect-v1，故而选择 Kinect-v2 作为家庭服务机器人的图像信息采集设备。通过获取 Kinect-v2 的 RGB 图像数据，对该数据经过图像预处理、人脸检测、人脸裁剪、特征提取、特征比对、阈值判断，最终确定图片中的人脸位置信息和对应的人物 ID。

2.2 亚洲人脸数据集制作

目前开放的人脸数据集大部分都是国外的数据集，如 Vggface2、CASIA-webface，虽然使用这些数据集训练出的网络模型可以在 LFW 测试集上取得很好的识别率，但是在亚洲人脸数据集或者在家庭服务机器人实际环境中的人脸识别效果非常差，所以建立一个亚洲人脸数据集是至关重要的。2018 年 5 月份，格林深瞳开放了 9.3 万 ID，280 w 张亚洲明

星数据，这个数据集是目前开放最大的亚洲人脸数据集，对学术界和工业界都有很大的帮助。通过统计数据集，发现该数据都是网络上的明星数据，缺少现实环境中的人脸数据，而且样本分布极不均衡，最多的一个 ID 包含人脸图片达到 3 000 张以上，最少的只有 2 张，样本不均衡在一定程度上会影响到最终的准确率。本文在亚洲明星数据集的基础上，结合通过使用家庭服务机器人采集的含有人脸的图片，构成了多场景、多角度的人脸数据集。同时对自建人脸数据集做了样本均衡处理，删除了部分人脸图片较少的 ID 和通过数据增强算法，包括加噪声等相关处理，将人脸图像较少的 ID 进行扩充。最终建立了一个 10 w 个 ID 的人脸数据集包含 350 w 张图片，该数据集简称 Ourdatas。人脸图像采集如图 7 所示。

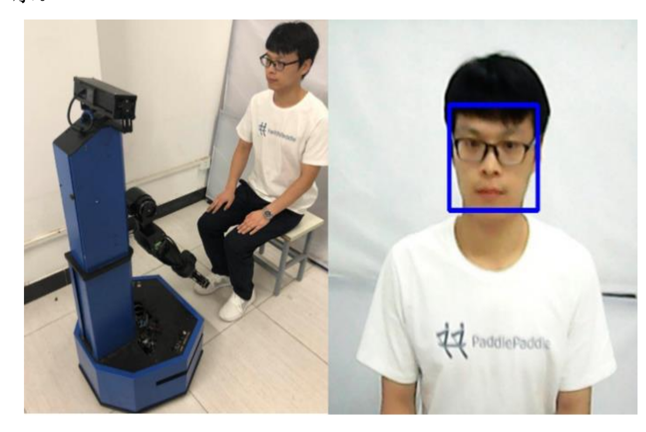


图 7　人脸图像采集

Fig. 7　Face image acquisition

2.3　实验结果

为了验证本文所提算法，在 Ubuntu16.04、GTX1080 显卡、TensorFlow 深度学习框架下，对 CASIA-webface 数据集和 Ourdatas 数据集进行训练，训练参数如表 1 所示。

表 1　训练参数

Table 1　Training parameters

参数	参数释义
--image_size 160	人脸像素
--model Inception_ResNet_V1	网络结构
--optimizer ADAM	优化器类型
--batch_size 90	图像批处理
--keep_probability 0.4	Dropout 参数
--random_flip	图像增强方法

其中：--image_size 是人脸图片归一化大小，本文中人脸定位后，裁剪人脸图片并归一化大小为 160×160 像素；--model Inception_ResNet_V1 为网络特征提取结构。--optimizer 为优化器，本文选择 ADAM，该优化器可以根据训练数据迭代地更新神经网络权重；--batch_size 为一次批处理读取 90 张人脸图像；--keep_probability 为 dropout 参数，可以减少过拟合，本文参数为 0.4，即每次迭代冻结 40%的神经元参数。--random_flip 是图像预处理操作。

2.3.1 通用测试集结果对比

本文的测试集为 LFW 数据集（包含 5 749 个 ID 及 13 233 张含有人脸的图像）和家庭服务机器人采集数据及相关竞赛现场数据（包含 2 400 个 ID 及 9 137 张含有人脸的图像，以下简称 OurFace）。训练集分别采用 CASIA-webface 公开数据集和自建数据集 Ourdatas。识别率的验证方法采用十折交叉验证，该测试方法是将待验证的数据集分成 10 份，轮流选择其中 9 份作为训练数据、1 份作为测试数据进行实验。实验结果如列表 2 和 3 所示。

表 2　CASIA-webface 数据集结果对比

Table 2　Comparison of CASIA-webface dataset results

算法	训练数据集	测试数据集	准确率
原始网络	CASIA	LFW	99.05%
改进网络	CASIA	LFW	99.22%

表 3　Ourdatas 数据集结果对比

Table 3　Comparison of our datas dataset results

算法	训练数据集	测试数据集	准确率
原始网络	Ourdatas	OurFace	97.78%
改进网络	Ourdatas	OurFace	99.10%

从表 2 和 3 中可以得出，以 CASIA-webface 和 Ourdatas 作为训练集，LFW 和 OurFace 作为测试集，本文的实验结果均高于原始网络。尤其是在亚洲人脸测试集上的准确率，较原始网络结构有很大的提升。

2.3.2 RoboCup 机器人世界杯现场测试结果

图 8 为 2018 年 RoboCup 机器人世界杯中国赛现场测试图。该世界杯包含家庭服务机器人多人辨识项目，项目的目的是测试家庭服务机器人在陌生环境中自主的识别目标人，标记出目标人的位置和 ID 及陌生人的位置信息。机器人通过语音交互和 Kinect v2 记忆指定的目标人，随后目标人进入到一个 5~10 人的小规模人群中，人群中每人姿势不定，有坐着、有站着。机器人需要找到人群的位置，并准确地识别目标人和陌生人。



图 8　2018 年 RoboCup@home 多人辨识结果

Fig. 8　Robocup@home multi-person recognition result at 2018

从图 8 可以看到，本文算法可以准确地识别目标人，图中用 Operator 标记，其他人标记为陌生人。在 RoboCup 机器人世界杯大赛现场，由于光线、角度、人脸分辨率等因素，FaceNet 官方的模型和传统算法误识别非常高，把陌生人误识别为目标人导致识别错误。本文的网络结构模型训练中加入了多角度不同光线下的数据集，同时也作了人脸对齐的相关处理，可以有效地减少光线和角度的影响，提升准确率。在 RoboCup 机器人世界杯上的出色表现，充分验证了本文算法的识别可靠性。本文收集了近三年 RoboCup 机器人世界杯现场人脸识别测试集，实验对比结果如表 4 所示。

表 4　识别率统计结果

Table 4　Statistical results of recognition rate

算法	测试集	正确识别	识别率
传统算法	30	18	60.0%
官方模型	30	22	73.3%
本文算法	30	29	96.7%

2.3.3 不同人脸误识别测试

官方原始的网络结构训练出的模型，虽然在 LFW 测试集上取得不错的准确率，但是实测中对中国人识别效果较差，不同的人脸特征向量也很相近，不能很好地做到类内距离近、类间距离远，经常出现不同的人判断为同一个。本文提出的改进的网络结构和建立的数据集训练出的特征提取模型，可以很好地区分相同的人和不同的人，相同的人计算余弦距离(归一化到 0~1，越接近 1 表示两个人脸越相近)得分很高，不同的人得分很低，通过匹配测试数据集，得出合理的距离阈值。本文中设置余弦距离的阈值为 0.8，高于 0.8 的判断为同一个人，低于 0.8 的判断为不同的人。图 10~11 中含有人脸的测试图片是两个不同的人。可以看出，本文的实验结果可以准确地判断出两个人脸是不同的人，官方模型判断两个人脸为同一个人，导致比对错误。实验结果如图 9 和 10 所示。同时本文测试 350 组不同人脸，判断算法把不同的人脸误识别为同一个人的误识别率，实验结果如表 5 所示。

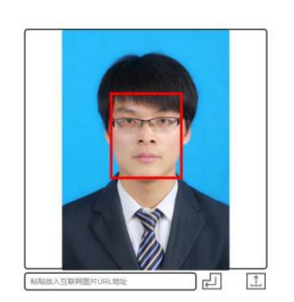
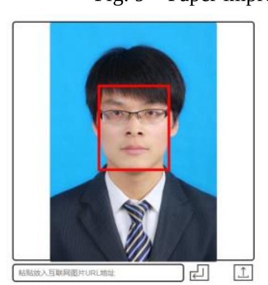





图 9　本文改进网络人脸比对结果

Fig. 9　Paper improves result of face comparison



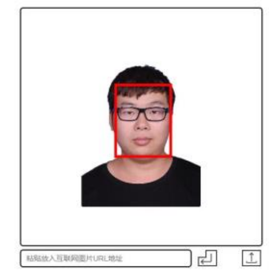



图 10　原始网络人脸比对结果

Fig. 10　Face comparison result on original network

表 5　误识别率统计结果

Table 5　Statistical results of false recognition rate

算法	测试集个数	误识别个数	误识别率
原始网络	350	43	12.28%
改进网络	350	12	3.43%

3　结束语

本文提出了改进的 Inception-ResNet-V1 的人脸特征提取网络结构，制作了亚洲人脸训练集和测试集。实验结果证明本文改进的网络结构和亚洲人脸数据集训练、超参数调节等相关工作，在 LFW 测试集和 OurFace 测试集及家庭服务机器人实际环境中得到充分的验证，实验结果均优于 FaceNet 原始网络。本文所改进的网络结构已经在北京信息科技大学 Sun@Home 家庭服务机器人应用，在国内多个服务机器人大赛中取得优异的成绩，充分验证了本文人脸识别算法的有效性。

参考文献：

[1] 徐圆圆，孔令富，高胜男. 实现家庭服务机器人中文指令解析问题研究 [J/OL]. 计算机应用研究，2019，36 (1): [2018-01-10]. (Xu Yuanyuan, Kong Lingfu, Gao Shengnan. Study on Chinese instruction parser issues for home service robot [J/OL]. Application Research of Computers, 2019, 36 (1): [2018-01-10].)

[2] 肖冰，王映辉. 人脸识别研究综述 [J]. 计算机应用研究，2005，22 (8): 1-5. (Xiao Bing, Wang Yinghui. Survey of human face recognition [J]. Application Research of Computers, 2005, 22 (8): 1-5.)

[3] 胡小荣. 复杂场景中家庭服务机器人目标识别与人脸识别的研究和实现 [D]. 合肥：合肥工业大学，2011. (Hu Xiaorong. Research and implementation of family service robot target recognition and face recognition in complex scene [D]. Hefei: Hefei Polytechnic University, 2011.)

[4] 方晨. 智能机器人“可佳” [J]. 科学世界，2013, 16 (11): 50-55. (Fang Chen. Intelligent robot"Jiajia" [J]. Scientific World, 2013, 16 (11): 50-55.)

[5] 张奇志，周亚丽. 北京信息科技大学——Sun@Home 家庭服务机器人 [J]. 机器人技术与应用，2014，36 (3): 49-50. (Zhang Qizhi, Zhou Yali. Sun@Home home home service robot, Beijing University of Information Technology [J] , Robot Technology and Application, 2014, 36 (3): 49-50.)

[6] Chan H, Bledose W. A man-machine facial recognition system: some preliminary result [R]. California: Palo Alto,Panoramic Research Inc，1965.

[7] Turk M, Pentland A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience, 1991, 3 (1): 71-86.

[8] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2002, 19 (7): 711-720.

[9] 彭辉，张长. 基于 K-L 变换的人脸自动识别方法 [J]. 清华大学学报 ：自然科学版 ，1997，37 (3): 67-70. (Peng Hui, Zhang Chang. Automatic face recognition based on K-L transform [J]. Journal of Tsinghua University :Natural Science Edition , 1997, 37 (3): 67-70.)

[10] Lu Xiaoguang, Jain A K, Colbry D. Matching 2. 5D face scans to 3D Models [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2005, 28 (1): 31-43.

[11] Taigman Y, Yang Ming, Ranzato M, *et al.* DeepFace: closing the gap to human-level performance in face verification [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S. l.] : IEEE Press, 2014: 1701-1708.

[12] Sun Yi, Wang Xiaogang, Tang Xiaoou. Deep learning face representation from predicting 10, 000 classes [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S. l.] : IEEE Press, 2014: 1891-1898.

[13] Sun Yi, Chen Yuheng, Wang Xiaogang, *et al.* Deep learning face representation by joint identification-verification [C]// Proc of International Conference on Neural Information Processing Systems. [S. l.] : MIT Press, 2014: 1988-1996.

[14] Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S. l.] : IEEE Press 2015: 815-823.

[15] Liu Weiyang, Wen Yandong, Yu Zhiding, *et al.* SphereFace: deep hypersphere embedding for face recognition [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S. l.] : IEEE Press, 2017: 6738-6746.

[16] Wang Hao, Wang Yitong, Zhou Zheng, *et al.* CosFace: large margin

cosine loss for deep face recognition [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S. 1.] : IEEE Press, 2018: 5265-5274.

[17] 陈耀丹, 王连明. 基于卷积神经网络的人脸识别方法 [J]. 东北师大学报 :自然科学版 , 2016, 48 (2): 70-76. (Chen Yaodan, Wang Lianming. Face recognition based on convolutional neural network [J]. Journal of Northeast Normal University :Natural Science Edition , 2016, 48 (2): 70-76.)

[18] Szegedy C, Ioffe S, Vanhoucke V, *et al*. Inception-v4, Inception-ResNet and the impact of residual connections on learning [C]// Proc of IEEE on Conference on Computer Vision and Pattern Recognition. [S. l.] : IEEE Press, 2016: 4278-4284.

[19] Szegedy C, Liu Wei, Jia Yangqing, *et al*. Going deeper with convolutions [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S. l.] : IEEE Press, 2014: 01-09.

[20] 杨文璐, 郭明. 基于 Kinect 的实时人脸识别系统 [J]. 计算机应用与软件, 2014, 35 (5): 64-67. (Yang Wenlu, Guo ming. A real-time face recognition system based on Kinect [J]. Computer Applications and Software, 2014, 35 (5): 64-67.)